DOCUMENT RESUME

ED 426 725 JC 990 052

AUTHOR	Lindner, Parker
TITLE	Assessment Tools for Distance Learning: A Review of the Literature.
INSTITUTION	Washington State Board for Community and Technical Colleges, Olympia.
SPONS AGENCY	North Seattle Community Coll., WA.
PUB DATE	1998-06-00
NOTE	34p.
PUB TYPE	Information Analyses (070) -- Reports - Descriptive (141)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Classroom Techniques; College Faculty; *College Outcomes Assessment; *Community Colleges; *Distance Education; *Educational Assessment; Evaluation Methods; Two Year Colleges
IDENTIFIERS	Washington Community and Technical Colleges

ABSTRACT

This literature review, created for the Washington State Community Colleges, provides a resource for faculty to use in selecting and constructing evaluation methods and assessment instruments for distance education. It was also devised to provide a starting place and tool kit for faculty to use in developing courseware and instruction in other than traditional classroom methods. The review addressed the following questions: (1) What studies compare distance learning classes to campus-based classes, and what tools do they use to do so; (2) What are the most current methods for assessing outcomes for each method of distance education; and (3) What studies suggest specific support systems that maximize successful teaching-learning interactions in distance education, and what are those suggested support systems? The review focused on community colleges and higher education only, eliminating most "training" studies, concentrating on controlled studies with measurable outcomes, avoiding anecdotal and "how to" studies, and including only recent (within the past 10 years) research. (Contains 49 citations.) (SL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 426 725

Assessment Tools For Distance Learning: A Review of the Literature

Funded through the auspices of the
Washington State Board for Community & Technical Colleges
and
North Seattle Community College
1997-98

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
P. Lindner
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Project Team

Parker Lindner, Project Director, Author

Tom Drummond
Shoshanna Porter, Reader/Analysts

Bill Moore, Project Advisor

Johnetta Moore
Zoe Holbrooks, Researchers

JC 990 052

BEST COPY AVAILABLE

Table of Contents

INTRODUCTION 3

PROJECT BACKGROUND 3
METHODOLOGY 5
DEFINITION OF TERMS 6
General Definition: 6
Correspondence 6
Interactive Video (ITV) 7
Telecourses 7
Online or Modem Based 7

GENERAL FINDINGS & OPENING DISCUSSION 8

TOOLS AND TECHNIQUES 11

REVIEW OF QUANTITATIVE METHODOLOGIES 11
Advantages and Disadvantages of Quantitative Measures 13
REVIEW OF QUALITATIVE METHODOLOGIES 13

QUESTION ONE 15

CONCERNS AND DISADVANTAGES 16

QUESTION TWO 17

CAN ALL METHODS OF DISTANCE EDUCATION BE EVALUATED BY THE SAME STANDARDS, STATISTICS AND MEASUREMENTS? 17
WHAT STANDARDS, STATISTICS AND MEASURES ARE APPROPRIATE FOR EVALUATING DIFFERENT DISTANCE METHODS? 18
Interactive Video 18
Telecourses 19
Online, Computer Mediated Instruction 20

QUESTION 3 21

GENERAL AND GROUPED OBSERVATIONS 21
Administration 22
Course Preparation and Instructional Design 22
Communication Techniques 23
Test Instruments and Continuous Assessment 24
Training of Assistants 24

SUMMARY AND CONCLUSIONS 27

Appendix A
Search Procedures Used
Appendix B
Reviewer's Forms
Appendix C
Telecourse Evaluation Questionnaire
Appendix D
Database Organization

Introduction

Purpose of project:
This project was created for the Washington Community College community. Its purpose was to provide a framework of understanding from which faculty could choose and construct evaluation methods and assessment instruments for distance education. It was devised to help instructors take advantage of the many studies and findings available in the field, rather than require each individual to start at the beginning.

Institutions are rushing to offer courses through technologies they have recently acquired, yet the lessons learned from the pioneers in distance education have not been widely disseminated. Though distance education has been expanding rapidly, only a small percentage of teaching faculty have taught courses in this way, or personally experienced distance based instruction.

It is hoped that this project will provide a resource bank, starting place and tool kit of assessment techniques that faculty can apply as they develop courseware and instruction in other than traditional classroom methods.

Project Background

In Spring 1997, the Assessment Committee at North Seattle Community College circulated a request for proposals listing a broad number of questions in the area of distance learning. The committee wanted to encourage faculty to test the outcomes of various distance methodologies.

Distance education, although well established as a niche within instruction, was new to most faculty members. There was need to establish a grounding somewhere.

This project was devised to provide the college with a starting place by engaging in a literature review regarding assessment methodologies in distance education. The proposal offered to review existing literature to seek answers to the following questions:

What kinds of controlled studies and data sets are appropriate for comparing distance delivered classes to classes taught in the classroom?

Can all methods of distance education be evaluated by the same standards, statistics and measurements?

In the community college environment, what interventions and support systems are required to maximize the potential for students to succeed in courses offered via distance education?

In accepting the proposal, the committee asked that the report:

- include details and specifics to help the committee understand the research methods, evaluation techniques and outcomes to be summarized in the report
- include advantages and disadvantages of the methods, techniques and measures discussed.

Bill Moore, assessment coordinator for the Washington State Board for Community and Technical Colleges was also interested in providing this starting place for other faculty members and institutions and in the Washington State Community College system. Moore offered to provide additional funding to leverage the original proposal into a larger and more inclusive project. As a result of this, the project was able to include additional reviewers, a librarian and professional researcher. The project team was comprised of B. Parker Lindner, Distance Learning Specialist, North Seattle Community College (Project Director), Tom Drummond, Child and Family instructor, North Seattle Community College and Shoshanna Porter, Chair, Electronics Department, Clark College (Reader/Analysts), Bill Moore, Assessment Coordinator Washington State Board for Community and Technical Colleges, (Advisor) and Johnetta Moore, and Zoe Holbrooks, University of Washington Graduate School of Library and Information Science (Research and Data Base).

Methodology

To initiate the research and assist the search process, a literature search work plan was devised. The questions were reframed to limit the search parameters. The librarians were to find articles and projects addressing the following:

- What studies compare distance learning classes to campus based classes? What tools do they use to do so?
- What are the most current methods for assessing outcomes for each method of distance education? (Correspondence, telecourse, interactive video, online/web based)
- What studies suggest specific support systems that maximize faculty's ability to teach and/or students' ability to succeed in distance classes? What support systems are recommended?

Because there is so much material on the subject, and in order to make the project manageable, additional search and selection parameters were agreed upon:

- Focus on Community Colleges and Higher Education (eliminate k-12).
- Use studies from 'training' only when they specifically assess outcomes.
- Focus on controlled studies with measurable outcomes.
- No anecdotal evidence unless it is part of a large scale assessment project such that anecdotal evidence could be quantified.[1]
- Focus on assessment, not "How To"
- Recent studies only, (last 10 years) unless seminal.

Over 100 hours of labor were devoted to reviewing key data bases, listservs and web sites (see appendix A for list). The search unearthed over 300 abstracts and URL's that appeared to answer at least one of the research questions. An organizational system and database were created using Microsoft Access 2.0® software to track selected documents and citations, and manage the document review.

After the researchers' initial work, article descriptions and abstracts were parsed to the team members who identified which of them should be pulled, found or purchased for review.

The materials were retrieved through the Washington State, University of Washington and Oregon State University libraries.

Evaluation forms were developed to provide a consistent set of analysis questions for the readers to use in reviewing the materials. (Appendix B)

Finally, the readers' comments were put together for this report. It should be noted that even among readers there were differences of interpretation of significance in findings and in choices of articles that might shed light on the project questions. Therefore not all the articles were read and analyzed by all of the readers. However, all readers have had a chance to review this report before publication.

Definition of Terms

For the purpose of this report, it will be important to agree on consistent terms to identify distance education modalities. As technological tools become more prevalent, formerly distinct delivery systems are merging and converging. Faculty are taking advantage of whatever tools serve their needs for demonstration, interaction, and reaching new students. In reviewing the literature, however, we noted that the distinctions are not always drawn. For example, some reports use the term 'distance learning' to mean only interactive video or telecourses. Others use the term telecourse but they are really referring to interactive video or cable delivered live classes. "Online" can mean anything from email to the WWW and the class can still be live and on campus.

These are the definitions to be used for the purposes of this report:

General Definition:

Distance Education is a planned teaching/learning experience where the teacher and learner are separated for all or part of the educational experience. Distance Education designs learner interaction and certification of learning using a wide spectrum of delivery technologies.

Correspondence

Correspondence classes, the earliest forms of distance education, were often used as the comparison or control group to assess the outcomes of new, more technologically rich distance systems. Correspondence classes are based primarily on print interactions for delivering materials, assessment and student-to-teacher communication. Of course, the telephone also plays a role in correspondence study.

Interactive Video (ITV)
In this paper, ITV will be defined as instruction offered in real time at multiple sites. The capability available through the video lines in the new Washington State K-20 system is ITV. In the documents read for this report, ITV does not always denote two way video but it always includes at least one way video (the instructor teaching live), with cameras and 2 way audio for student interaction.

ITV is distinguished from telecourses because the action is synchronous -- that is, students and instructor are working in real time. ITV delivery systems may vary. Some ITV is offered via compressed video that is hardwired or microwaved while other institutions might use satellites to deliver the content over a broader area (called the footprint). For the purposes of this report these delivery differentials have been ignored unless relevant to study outcomes.

Telecourses
Telecourses are defined as pre-produced, video tape based courseware. Telecourses are rarely tape only but are packaged with instructional materials including study guides, faculty handbooks, suggested textbooks or readings and test banks.

Online or Modem Based
Online classes rely upon computer based interactive technologies for content delivery, faculty-student and student-student interaction. E-mail or computer conferencing tools are the cornerstone of the class interactivity. More sophisticated systems would provide for such options as online testing, animated simulations or use of audio files. Generally, the term online course is used to define anything from campus based classes that use email for out of class discussion, to courses that make extensive and exclusive use of the world wide web for content delivery, interaction and research. This report uses the term to refer to online classes that are asynchronous -- where the instructor and learner are not working at the same time or place.

General Findings & Opening Discussion

Distance education has been the subject of research ever since correspondence courses became popular at the beginning of the nineteenth century. Until then, education was predicated on a centralized repository of knowledge where information could be presented to the student. The classroom provided the main opportunity to synthesize that information, where the instructor managed and guided the discussion. Many instructors who are comfortable in this milieu fear that distance education might eliminate the dynamic exchange between students and their faculty as well as their peers.

However, findings of the research easily dispute that notion. Recently, Thomas L. Russell of North Carolina State University posted a page on the world wide web[2] that chronicles 248 studies in which the compared learning outcomes of students in various distance modalities could show no statistically significant differences from classroom instruction. Even so, skeptical faculty want to know how these conclusions were reached and under what conditions. [3]

This review made every effort to seek out **scientifically constructed** and empirically replicable research projects. In reviewing articles, all of the readers noted that many of the so-called research papers were really descriptions of programs or lessons learned. We have tried to clearly distinguish such subjective evaluations while at the same time retaining articles in the database if they contained useful insights.

It should be noted however, that empirically provable and replicable research techniques may not be the most appropriate or useful tool for the instructor/practitioner. Controlled statistical studies are generally the domain of researchers and graduate students. It is unlikely that the average community college instructor would want to utilize such techniques in their day-to-day approach to instruction. Often community college classes contain fewer than the N=30 requirement for statistically useful distributions, limiting the statistical value of studying only one class. Ideally, studies should be conducted over several classes and/or over time in order to be quantitatively validated.

Even without replicable controlled research, educators still need useful tools to assess programs and evaluate students in a distance learning context. Therefore, this report has also sought out other means of identifying how and whether learning has taken place.

As educational reform pushes us to reconsider what outcomes we actually assess and grade, the test scores and mean grades of many of these studies must be put in perspective. This reconsideration is especially important in light of findings such as Cheng, et al (1991)[4] that show distance learner's achievements more equivalent when assessed by methods emphasizing competencies over content. In other words, distance learning challenges us to evaluate student's actual knowledge construction and integration rather than simple absorption of facts.

In addition, many of the studies rely upon either students' or faculty members' subjective impressions or levels of satisfaction. Scientists may find these findings specious while others argue that such insights can provide valuable information.

The readers have also noted that the research often fails to distinguish the assessment of learning ***systems*** from the assessment of student/learning ***outcomes.*** In this regard, several of the articles read for this review advised that assessment involves different purposes and approaches depending on the intended audience.

Researchers who are trying to prove or disprove the efficacy of distance education generally use grades, attitude surveys and observational data to suggest the quality or significance of learning or instruction in distance education. Few of these studies actually assess learning outcomes. No matter what the modality, assessment should conform to some standard.[5]

One of the most useful and realistic studies in this review was one undertaken by the RAND Corporation at the behest of Annenberg/CPB to attempt to evaluate telecourses and their outcomes. Even a reliable and highly respected research operation such as RAND had great difficulties creating a controlled and replicable study in the context of the autonomous culture of higher education (Shavelson et al, 1986).[6] The report suggests that the most critical factor in telecourse evaluation was how the course materials were implemented. The authors note

> " Implementation issues intruded so often as to make them central to the interpretation of outcomes. Future telecourse research should place special attention on implementation, not only because it influences student outcomes but also because it illumines a variety of important general issues concerning the future of "distance teaching" in American higher education. This requires on-site monitoring of all aspects of telecourse implementation.
>
> ...just because telecourses seem more like predetermined packages than traditional courses does not mean that instructors will teach them that way"[7]

The RAND report suggests a "grand lesson" -- that one should expect extreme variation and unreliability in operationalizing research designs in the American system of higher education. In order to overcome these variations they suggest that any research must over sample the population, provide plenty of lead time, include faculty in the research design, and provide monetary incentives for faculty participation.

Audiences

Some papers emphasized that faculty, administrators and students will each have their own questions regarding these delivery systems. (University of Wisconsin-Madison, 1987[8]; Shavelsen, *et al.* 1991[9])

Administrators are concerned with cost/benefit analyses, and the political implications of distance education. They want to know, "does it make sense structurally and financially as expected? Did students achieve goals? Are programs financially sound"? Students want to know how the distance education experience compares to traditional delivery, how effective it is, and whether it is user friendly and attractive to learners.

Faculty will be most concerned about how well distance techniques achieve learning outcomes and how much additional effort or support might be needed to achieve comparable outcomes through various distance learning methods. Faculty are interested in the factors for success and what tools techniques and training will be needed to achieve them. Faulty want tools for instructional preparation, learning assessment and course evaluation. In writing this document it is tempting to report the findings and outcomes of the studies read. However, the objective of this project was to identify useful tools. Readers are encouraged to read the articles or undertake further studies regarding specific outcomes.

Tools and Techniques

Review of Quantitative Methodologies

Since the intent of this research was to ferret out methods of analysis regarding distance education, it will be useful to review the most frequently used analytical methods and empirical research techniques. In other words, what is measured and how are those measurements evaluated?

In large part the data sets utilized for distance education research start with elemental **descriptive statistics**. Researchers typically process data sets comprised of test scores, grade point and dropout or retention rates. Often these raw data are analyzed by calculating the mean and standard deviations, then looking for relationships with other measures such as demographics, time on task, specific methodologies and student attitudes.

Attitudes and satisfaction rates as well as perceptions were most often quantified by utilizing the four or five point **"Likert scale"** with a measure from high to low. Of course, the quality and usefulness of such a scale is directly correlated to the quality of the questions asked. Very few of the reports read actually included their survey instruments.[10]

Most typical evaluative measures were variance and covariance analysis where scores such as pre and post tests, grade point or student survey differentials were analyzed for their statistical significance.

Statistical significance means answering the question: was the difference between the groups or data sets measured significant because of the difference among the variables or could that difference have happened by chance alone? Mean scores of the samples are placed into a standard formula which determines a test statistic*(T value)*.[11] Then a probability table is consulted to determine the **probability** this outcome was the result of the variable being measured, or if it could have happened by chance. If the outcome is statistically significant, it means there is a high probability the outcome happened because of the variables measured.

Covariance tells whether there is a linear relationship between pairs of variables and whether that relationship is positive or negative. If the relationships are not linear (i.e. not directly causative) this technique should not be allowed to imply causation.

In complex situations it can be assumed that an outcome such as a test score is the result of many factors. **Regression** analyses are used to calculate the relative influence of those factors.

A disadvantage in studying distance delivered classes is always "selection bias." In other words, is there a natural bias if students *elected* to take a distance class? Some of the studies addressed this bias, others did not. For example, in Craft and Wagner, (1988)[12] data were compared between on-campus and off-campus ITV students. Covariance was used to sort out factors that could bring about differences such as demographics or instructional effectiveness. However, typically there was no statistical significance!

In Morehouse (1987)[13] the firm of QED created a meta-analysis of Minnesota's Technology Demonstration Program from 1983-1987. Using a "variety of analyses and testing procedures" (pp119-122). Researchers compared student achievement in ITV courses with achievement in the same courses delivered traditionally. No consistent, statistically significant difference was evident in any of the nearly 1000 grades and test scores analyzed.

Seamons (1987, p142)[14] studied "The Influence Of Teaching Style And Instructional Device Use On Student Satisfaction And Preference in Electronic Distance Electronic Methods" at Utah State University.
He used correlation/coefficient matrices and prediction models with multiple regression to identify relationships between teaching styles, student satisfaction[15] and performance. Student performance was measured by mean grades, and student satisfaction was inventoried with evaluation forms. Seamons also used quantifiable observational data such as library book checkout, A-V usage, length of papers, and assignments done on time.

Paul L. Beare (1989)[16] utilized a Chi squared procedure to look for significant differences among instructional group formats. He analyzed six different delivery methods with the dependent variable being the method of instructional delivery used and the independent variable being the percent correct on exams and course evaluations. The data showed that the instructional formats had little effect on student achievement or course evaluation. A drawback of using only this method of evaluation is that it does not provide any measure of teacher/student interaction or other implementation factors.

In "Evaluation of Student Outcomes In Distance Education," Feasley (1987)[17] provides good information on how to do validity/reliability testing on exam questions. His paper explains mechanisms of test construction such as indices of difficulty for questions and discrimination factors. The purpose of such analyses is to sort out how well students at various levels of performance can correctly answer a given question.

Advantages and Disadvantages of Quantitative Measures

Quantitative analysis, when well designed and carefully executed, can be useful for comparing numerical results such as test scores. However, such studies must be scrutinized. One must always pay careful attention to what numbers actually measure. It is particularly important in assessment of student learning to consider the difference between "outcomes" - the measure of whether a student possesses the skills and knowledge desired from the course material -- and "gain"- whether there was a gain in the student's knowledge. Simply comparing test scores without understanding or judging the tests themselves can create unreliable conclusions.

Review of Qualitative Methodologies

There have been many reports, papers and studies that looked for specific qualities in distance education. These qualities are often described using observational data, or they are identified, measured on a scale of satisfaction, and then turned into numerical data.

Many studies focused on perceptions and preferences. Although they may have been quantified via a Likert scale for comparison, ultimately such judgments are subjective and must be viewed in context of how the original question was stated and framed. Very few of these articles actually included the survey instruments themselves so they should be considered qualitative. This is not to devalue the research. For example, Jackman and Swan (1996)[18] looked for effective instructional models in a distance education classroom Rather than assess the students, the study looked at student preferences for instructional models applied within the distance education/ITV structure.

The Likert scale is the most common way of turning qualitative responses into numbers, asking for a measure of quality on a scale of one to four or one to five. As mentioned before, this method is only as useful as the consistency and clarity of the questions. Student satisfaction is generally measured in this way and a number of researchers have drawn on the work of Paul M Biner,[19] using or modifying the questionnaires he has refined.[20] Likert scales are frequently used in surveys to measure attitudes and perceptions.

Focus groups are another method of finding out how individuals feel about situations. Researchers often conduct such groups to gain a sense of attitudes. Focus groups are useful if carefully moderated and focused on specific objectives.

Many studies rely on instructor and student self evaluation and documentation. From this information, conclusions are drawn about the experience of or satisfaction with the learning process. In scientific research, such self reports would likely be excluded as too subjective and not empirical.

Researchers also relied on observational data as monitored by third parties such as graduate students, room monitors or teaching assistants. They would observe, for example, the amount and quality of interaction and engagement between students and teachers or student and students, comparing those interactions between the televised classroom and the 'live' situation or skills and performance based on observable behaviors

Feasley (1987) points out that if such analysis tools are to be used, the observers must be carefully trained When observational data is to be gathered, Feasley emphasizes observers of tasks as well as performers of those tasks should be given pre-prepared instructions that detail both the task to be done and the specific observations and data to be gathered.[21]

Erping Zhu's (1996)[22] fascinating and useful study "Meaning Negotiation, Knowledge Construction and Mentoring In A Distance Learning Course" analyzed the notes from electronic discussions and coded them into participation categories such as question, reflection, discussion and answer, creating an "interaction analysis." Using this technique, Zhu was able to create a measure of quality in knowledge construction using Vygotsky's theories of learning as a process of social negotiation and collaborative construction.

Readings also mentioned graduate and employer surveys, synthesis activities and portfolio assessment but the studies gave no details or clues to their development, structure, or the construction of rubrics.

Question One:

What kinds of controlled studies and data sets are appropriate for comparing distance delivered classes to classes taught in the classroom?

The most commonly used data for making classroom/distance comparisons are grades and test scores. When comparing groups of classes, mean scores are needed to create the test statistic. Usually the control group is the traditional classroom or the site where the instructor is "in person" (in the case of ITV). Occasionally score differences on specific knowledge or behavior tests are noted. Retention or completion rates are also of interest.

In Cheng (1991)[23] an on-campus control group of 25 students was compared to a class offered by computer conferencing. Pre and post tests, total semester scores, lab time, project assignment tools, outside class computer time and total time on task were all measured. The scores for these were analyzed via ANOVA and ANCOVA (analysis of variance and analysis of covariance) backed up with questionnaires and achievement tests. The distinction between ANOVA and ANCOVA was interesting because scores that were statistically significant in the ANOVA tests (t & p) became less so when correlated with age. While the study found no significant differences in final grades, it did find that in a section where participants worked jointly in study sessions there was a 90% completion rate. The study concluded that distance learners do better with project based activities and when given joint study opportunities. The author suggests this is because computer students spend less time with lecture materials and more time on activities. Correspondence students generally spend the most time on task. While this report thoroughly measured a number of variables, one disadvantage is that the reader has no way of knowing the qualities and internal tools included with the computer conferencing software used.

Paul Beare (1989) compared six delivery methods by measuring test scores of students in different sections and then seeking statistical significance via Chi Squared and ANOVA analysis. In written comments, students preferred live instruction -- particularly the opportunity to ask real time questions -- but this was not tallied in the study. The report suggests that students should be randomly assigned to the different media groups if possible. Otherwise the students would self select into classes based on their learning preferences.[24]

Whetzel, *et al.* (1996)[25] studied courses at the US Postal Service comparing satellite delivered to classroom delivered training. Researchers evaluated student attitudes and student reactions and assessed learning through pre and post tests using covariate analysis. In this study, the satellite learners showed

higher learning gains than the classroom group. According to the author, the research results were consistent with the Office of Technology Assessment's appraisal that satellite training can provide a consistency of learning gains within a diverse instructional environment. Even with the wide range of courses and pre-test skills of the students, the study was able to identify distinctions between satellite training skill outcomes and classroom skill outcomes and to identify possible variables that might affect student learning. These included on-site facilitation, interactions from the remote sites, the nature of the target population, instructional development and dissemination strategies. However, the organizational complexity of the postal service and range of courses offered, precluded true research design using a control group.

Jackman and Swan (1996) chose to measure student perceptions of instructional models used within the ITV classroom. Their survey instrument used a five point Likert scale to evaluate items related to instructional models,[26] the technical system and demographics. It also used the Cronbach Coefficient Alpha Test to determine reliability of sections of the instrument. Using this scale, Jackman was able to assess preferences for instructional types. Highest ranking was role playing, then simulation, jurisprudential, memorization, synectics and inquiry. Direct instruction was deemed the *least* effective model in the ITV classroom. The off-campus group placed more importance on the memorization and conceptual models than did the on-campus groups, and both groups placed high importance in active learning models.

Concerns and Disadvantages

Feasley (1987) points out the need for controls and consistency in test construction through known techniques such as establishing the validity of the question, the extent to which a test measures what it is supposed to measure and the reliability of the test. He asks, "is there consistency of student performance within the test over time?"[27]

RAND noted that pre-tests are a rarity and instructor-made exams often test on the textbook rather than the course content. The RAND report emphasized the importance of studying course implementation so that variations in approach (such as use of materials, testing, student characteristics, participation, number of in-person meetings, grading schemes) can all be taken into consideration.[28]

Question Two:

Can all methods of distance education be evaluated by the same standards, statistics and measurements?

Studies comparing courses and course delivery systems have found that modality may not be the most important indicator in course or class success. In "Variables Affecting Adult Learning In a Distance Learning Setting," researchers Fellani, Blackwood and Seamons(1987) concluded that students were more likely to choose a course by professor - than by technology. In "Evaluating Interactive Television: Methods, Findings and Issues" (1987) Morehouse demonstrated there are more differences between instructors than there are between modes of delivery.[29]

Instructional designer Jeff Hoffman[30] asserts that his "Learner Interaction Model" can be applied to any distance learning medium. The model he advocates elaborates on the 1989 work of Michael Moore, suggesting four areas of concern:

Learner-Content
The intellectual interaction where the student's understanding, perception and cognitive structures are changed.

Learner-Instructor
Where the instructor plays two roles, "human touch" and "director of learning"

Learner-Learner
Often overlooked, but where students expand and apply their knowledge of the lesson content.

Learner-Interface
The student's interaction with the delivery mechanism-- is it user friendly and transparent?

Cheng *et al,* (1991) compared three learning groups by mode of course delivery--straight correspondence, computer conferencing and classroom. On achievement tests using post-test scores, the correspondence group scored significantly higher than on-campus and computer conferencing groups (in that order). In the final grade, however, which included both tests and projects, there were no significant differences between the groups. This report suggests that since the post test was predominantly a low level recall exam, those students who were more familiar with the course lecture materials would do better in the examinations. However, where overall course performance involves more than recall of lecture material, "delivery systems did not have a significant effect on

understandings or competencies that could be shown through application of concepts."[31]

What standards, statistics and measures are appropriate for evaluating different distance methods?

Interactive Video

Focus groups were the basis for Ellen Wagner's 1993 article which identified programmatic analysis needed in developing interactive televised courses. These included technological reliability, institutional support, the adaptation of traditional support structures to the needs of distance learners and program planning including development of procedures to address student and faculty concerns.[32]

Paul Biner's early work is often cited and used as a guide for developing distance learning survey instruments. Biner (1994) argues for the value of assessing student satisfaction levels and he reiterates this belief in an updated article "Reassessing the Role of Student Attitudes in the Evaluations of Distance Education."[33] Biner's work used factor analyses to statistically identify common groups of items that, when grouped together, become the 7 key factors that comprise student satisfaction. The factors are:

- the instruction
- the technology
- course management
- at-site personnel
- promptness of material delivery
- support services
- out-of-class communication.

An advantage of Biner's work is that his first study (201 students) was validated a year later with 177 students when the study was repeated. A disadvantage is that factor analytic studies rely on judgment for identifying the factors. Also, it can be argued that satisfaction factors are not useful or appropriate indicators of knowledge outcomes.

Hackman and Walker (1994) compared students who were taking the same class but one group was on-site with the instructor and the other group was at the remote site. It used one way ANOVAS to study system design factors and communication norms in televised instruction(one way video and 2 way audio). The study, unlike many of the others, achieved a 100% response rate on

questionnaire completion and thus was a reliable indication of communications perceptions.[34]

Eagan and McCleary (1989) studied student outcomes measured by tests and retention, instructor effectiveness, learner receptivity, level of difficulty, amount of material covered, comparison to a conventional course and course design features.[35] The research followed progress of a three course sequence, introducing procedures and conditions over time. The study documented changes that were implemented based on feedback and assessment. Through this feedback, improvements were made in clarifying the role of the site facilitators, student evaluation procedures, provision of learning materials, instructors visitation of sites and use of visual media. Although the research is not replicable because the intervention procedures changed over time, the researchers learned about important success factors in this type of instruction. There were statistically significant differences when visual material and media were used to clarify concepts or demonstrate application of techniques. As instructors redesigned the course manual to closely correlate with content in each course section, learner's perceptions of organization significantly increased. Student feedback showed a large gain after on-campus instructors developed specific procedures to help off-site staff convey feedback to students . This correlates nicely to the learner interaction approach suggested by Hoffman.

Kendall and Oaks (1992) simply used a questionnaire to survey faculty perceptions about teaching over the Washington Higher Education Television System(WHETS). This instrument offered the faculty members an opportunity to consider what tools and techniques seemed to work. The study analyzed adaptations in course delivery, audio-visual aids and interaction with the students.[36]

Telecourses

The goal of the RAND report was to establish appropriate mechanisms for assessing telecourses. This is the report that struggled with the "strongly ingrained tradition of autonomy in American higher education." The project was specifically looking for evaluation designs for telecourses. Researchers experimented with variables such as verbal reasoning, achievement, background, study habits, and attitudes towards telecourses; then attempted to correlate these with the telecourse curriculum model and implementation data. During the study's three year cycle, RAND observed that telecourse implementation varied significantly from instructor to instructor. Researchers concluded they could not consistently compare the outcomes without also studying implementation factors such as number of meetings, use of video lessons vs. text materials and testing methodologies. The evaluators rejected most traditional research designs in favor of a pre and post test between two consecutively taught sections for the

same telecourse. The researchers analyzed student characteristics from each site group, interviewed site administrators and instructors and did pre and post tests for student competencies in the subject matter. Using this method, RAND found that gains in test scores were remarkably similar from one semester to another despite variation in telecourse implementation within and across sites and different examinations administered by the different site instructors.[37]

The RAND report strongly emphasizes the importance of measuring outcomes -- whether the student possesses the knowledge, skills and attitudes desired/required, rather than measuring 'gain' in knowledge. This report also asserted that telecourse research design should allow for student choice of course type because students choosing telecourses differ from those choosing traditional courses. If choice is not permitted, they believed, attrition would most likely be non random, making the results uninterpretable.[38]

Online, Computer Mediated Instruction

Erping Zhu (1996) studied the documentation from a course taught through computer conferencing using a software system called VAX notes. The researcher investigated how and whether learning was occurring within an online discussion group and class. Zhu was influenced by the theorist L. S. Vygotsky (see note 22) who studies the relationship of social processes and interactions to knowledge construction. Using Vygotsky's work as a model, Zhu was able to assess the learning process through content and role analysis. Working with instructors who were knowledgeable in instructional design, Zhu had three specific discussion components to study:

- the instructor's introductory questions and reading advice
- student generated weekly starter comments,
- student generated summary (wrapper) comments, participant comments and discussions.

By coding the discussion notes into specific categories, types, and directions of interactions, the author was able to identify the students' patterns of learning in a replicable manner.

Zhu used a coding mechanism that analyzed the class interactions by defining participation categories: question, reflection, discussion, comment and answer. In addition he observed vertical vs. horizontal interactions - which interactions showed the students learning from peer to peer (horizontal) compared to vertical interactions, up the chain to a higher authority. He also defined student and instructor roles such as contributor, wanderer, seeker, mentor. Using this framework, Zhu analyzed and measured quantity of interaction by volume of

content in the postings and quality of interaction through the nature of the interactions and roles as identified and coded.

Using this method Zhu was able to prove where participation rates were shared (not dominated by extroverts), and was able to give concrete information on the interactions through textual analysis of comments made and their content. The study was able:

- to show multiple perspectives as a rich context in construction of understanding
- to prove that when properly executed, electronic discussion allows students to discuss more than one thing at a time
- to demonstrate that conferencing software is a useful tool for reflection apprehension.

Question 3:

> **In the community college environment, what interventions and support systems are required to maximize the potential for students to success in courses offered via distance education?**

Students of distance learning have done many studies to answer this question. Thousands of instructors who care about the quality of their instruction have spent countless hours of observation and self reflection to answer it as well. The answers are no mystery. They are primarily the same answers required of any coordinated instructional effort, amplified by the extra need for clarity, feedback and systems necessary to achieve the goals for students who are learning at a distance.

Now that distance education is firmly entrenched as a delivery option for many community colleges, faculty and administrative training opportunities abound. It is beyond the scope of this report to adequately cover the field. Included here is a general summary of the interventions uncovered in this literature review.

General and Grouped Observations

In Wagner (1993), focus groups at the Western Cooperative for Educational Telecommunications[39] worked to articulate variables affecting distance education programs and student success within them. The groups identified as most significant

- needs assessment
- audience analysis
- instructional design (to include analysis, design and evaluation),
- well stated objectives written in performance based terms
- lesson planning and interactive instructional strategies. [40] [41]

Barry Willis (1992)[42], who pioneered distance instruction in Alaska in the 80's, identifies a set of process variables to ensure student learning, and programmatic success. These are:

- the promotion and facilitation of student to student interaction,
- having the students contribute information about their personal backgrounds such as cultural, social and geographical background and interests
- practice in using the technical delivery systems to overcome technology related obstacles and promote teacher-student communication.

Interventions rarely change with delivery system. Hardy and Olcott (1995) who studied classes delivered solely by audio teleconferencing, came to many of the same conclusions as other studies - that the evidence of comparative performance must be ... "related to interaction, technology coordination, support services and training."[43]

Administration

Distance Education has been shown to be most effective when student and administrative services are consolidated in support of distance instruction.

Craft and Wagner's report (1988) advocates for centralization of support functions such as materials duplication and distribution, grade reports, admissions records, troubleshooting technical problems, credit transfer, room scheduling. Researchers found that for ITV classes, daily courier service was also a success factor. This article calls the support staff the 'silent stars' in distance education' and states that administrators of distance education should not relinquish authority and decision-making to technical managers.[44]

Course Preparation and Instructional Design

"Insufficient on-campus administrative support with regard to time and materials for course redesign was the most significant deterrent to course success."[45] (Seasons) This is reinforced in Egan and McCleary which suggests the efficacy of complete pre-authored course manuals.

Several studies indicated the value of varied and appropriate instructional models. In Jackman and Swan (1996) the highest ranking instructional model was role playing. Other preferred models included simulation, jurisprudential, memorization, synectics and inquiry. Direct instruction was deemed the **least** effective model in the ITV classroom. The off campus group surveyed in this study placed more importance on the memorization and conceptual models of instruction than did the on-campus learners. Both groups placed high importance on active learning models.

In regard to telecourses, the RAND report called attention to instructor initiated strategies such as review sessions for off campus students prior to exams, use of the telecourse study guide, on-campus meetings with the instructor and supplementary lectures.

Cheng, *et al.* , (1991) suggests that joint study sessions have a significant effect on retention/completion rates for distant learners.

Studies by Hoffman, Winston, *et al.*, and others indicate that good illustrations and visuals enhance student performance and understanding.

Communication Techniques

Early work in correspondence studies has brought forth a set of interaction techniques considered good practice in working with distance students. These are articulated in correspondence study manuals cited by Tom Drummond such as Rowntree and Estabrooke (both 1981).[46]
among the ideas: Instructors interacting with students at a distance instructors should

- Use first names.
- Reference pride " you can be proud you..."
- Indicate the instructor's faith "I believe you can keep up the fine work"..
- Indicate the instructor's appreciation " Thanks Jean, you've brightened my day."
- Share the instructor's experience "During my 10 years of teaching I have not seen a better..."
- Reiterate instructor's willingness to help "I'm here to help you any way I can.."

A good list of feedback techniques can be found in Feasley, p23.

Research stressed the importance of training instructors in the use of "immediate and socially present behaviors" which include calling on students at a distance

by name, inviting and recognizing their participation and disclosing personal data. (Hackman & Walker, 1994)[47] The instructor's training as a whole was considered significant to student perceptions of the distance course and their instructor's effectiveness.

Test Instruments and Continuous Assessment

Feasley (1987) gives specific information on test construction with validity testing and reliability testing. The RAND report emphasized the difficulty of comparing courseware because test instruments are so inconsistent between faculty members. Cheng, *et al.* (1991) suggests that project based assessments are more likely to capture the learning outcomes of distance learners. Because distance education invites extra scrutiny, it is critical that faculty hone their skills in test construction and instructional design (Feasley & others). In addition, because there is no opportunity to do traditional observation, continuous assessment in the form of asking questions and assigning tasks are critical to maintaining contact with student performance in a distance learning context. (Feasley)

Training of Assistants (site people, test graders)
Often distance education is a team effort. Online instruction requires students to become oriented to the tools they will be using. Someone must troubleshoot technical obstacles. Telecourses require tape and materials distribution. ITV may include technicians as well as site facilitators. Eagan and McCleary (1989) stress the importance of adequate training and carefully constructed roles for site facilitators since they act as a bridge between students and instructors. Adequate orientation to the technology is also critical to students' real and perceived success.[48]

The articles reviewed in this project represent a broad variety of attempts to identify and measure key aspects of distance learning. None can stand alone as providing either answers to the questions we ask nor models for our own assessment efforts. As in all educational endeavors, proof of learning efficacy eludes an absolute determination.

Yet assessment remains the key to attaining effective learning. It is after all, how students see what they are supposed to do and how teachers, in turn, find out whether they are effective at what they do. Distance learning presents even more challenges to effective assessment than traditional face-to-face classroom teaching, because immediacy, visual contact, personal contact, and complex relationships are hampered. Additional checks and balances must be in place to ensure the quality of teaching and learning.

Many of the studies reviewed here provided insight into the factors that must be considered in designing, implementing, and checking the results of distance

learning, no matter which modality is used. The table presented below represents the components that deserve attention when designing assessment efforts. Information can be gathered about each input, each process, and each output.

If all are addressed, important information can allay the concerns of all audiences who care that distance education works well for students who cannot attend a campus course. One may not be able to prove that an ITV history class is as effective as a traditional one, but one can ensure that every possible attempt has been made to make it a good experience and that means are in place to uncover elements to modify or correct in the future.

Inputs	Processes	Outputs
students	**instruction**	
prerequisite skills	2-way communication established: students introduced to talking freely	**student reactions to training**
expectations	information gathered about each student	
instructors	practice using technical tools	**amount of student learning**
proficiency with technical tools	content related examples • modified to fit needs of learners • predominate instruction time	knowledge reasoning skills dispositions/attitudes
cultural familiarity with students		performances of understanding
application of previous formative evaluation information		
visit sites	**instructor behaviors**	**behavioral change on the job**
site coordinators	promptness online	
ability to troubleshoot	natural, clear delivery	
discussion leadership skills	spontaneity frequently change pace	
possess content knowledge	draw out participation use participant names	**organizational results**
assess student learning	give short summaries	
course organization	control verbal traffic	
performance objectives for course	share decisions democratically	
each session performance objective	create shared space and history	
content related examples	model appropriate behavior	
graphic design	clarify common definitions of terms	
active learning	set and check pacing	
student preparation materials for each session	conduct classroom learning assessments	
support services	**site facilitators**	
advising	seek out students needing help	
library resources	provide feedback on performances	
technical services	**technical**	
maximum clear connections	"up" time	
email/computer discussion services	email/computer utilization	

Summary and Conclusions

The question is not which medium works best but what is effective instruction.

(Whittington 1987)[49]

The intent of this paper was to set up a framework of understanding of the tools and techniques used to evaluate distance education. It set out to answer three questions that were quite global in nature.

What the readers found was that although there have been many studies that used empirical mathematical techniques, the value of these numbers is limited to one's perception of the elements they chose to measure. It is most illuminating that despite Thomas Russell's WWW posting of 248 studies that found no statistical significance between distance and classroom instruction, skepticism remains about this mode of instruction. This suggests that while innumerable studies have been conducted, few have sought to answer the right questions.

In many fields of endeavor besides education, technology has managed to become the subject rather than the object of discussion. In this review we have seen that technologies or modalities of instruction are less important than the quality and design of the instruction and communication they convey.

We have also seen that the measurement tools and data sets for statistical analysis have been, for the most part, the traditional measures of grade point, test outcomes and attitudes. In correlating these measures, certain rules and truths have been established:

That distance learning engenders more scrutiny than classroom instruction.

That, in order to stand up to that scrutiny, additional program planning is required, especially in instructional development and procedures to address student and faculty concerns.

That test procedures and measures to evaluate student outcomes should be scientifically designed.

That the quality and quantity of interaction has a high correlation to the success of distance learning and that active learning instructional models work better in the virtual classroom than didactic instruction.

That group process, collaborative learning and joint study sessions are successful instructional models for this learning context.

That institutional support of this interaction must occur on a variety of levels to ensure the reliability of the technology used, the ability of the faculty to redesign courses to take advantage of the technologies, and the friendliness of the medium from the learner's perspective.

Instructional evaluation should be predicated upon a specific set of values, expectations, goals and objectives. We found that although there have been volumes written and researched about distance learning, few studies actually illuminated our understanding of how knowledge is constructed or where, when and how learning takes place.

From the literature of assessment, we do know that getting and using formative data improves instruction. In the research on distance education, where formative data was collected and used to construct changes in instruction there were significant improvements in outcome. Egan and McCleary (1989) found this occurred when improvements were made in encouragement of and access to student feedback, selection of content related examples, clarification of expectations and materials preparation. Small surprise.

The RAND report had it right - that implementation is the critical factor in distance education. Instructional design must be seen to include much more than the physical look and feel of the materials but to include the instructional models utilized, clarity and design of course and lesson objectives and the incorporation of assessment techniques.

Further investigation should ask: What additional and alternative observations and interventions can be used to assess learning and student outcomes in the context of distance education? How should distance learning courses and faculty be observed, measured and evaluated to assure academic standards?

Distance learning is no longer a tangential sideline to community college education. It is one among the many ways we make education available and accessible to our students, and it is not likely to disappear. Further inquiry into whether the mechanism of delivery has value is not necessary. It is more significant that we change the focus to assessment of learning.

Notes and Citations

[1]This parameter was the subject of debate among the committee. Drummond and Moore asserted that anecdotal evidence is necessary to make qualitative assessments. Because there was so much material the team agreed to revisit this question if deemed necessary after reading the literature gleaned from this process.

[2] Available at http://wwww.oit.cmich.edu/phenom.htm

[3] Russell states: "While this documentation speaks volumes about the futility of these studies, it also acknowledges the fact that the questions about the comparative impacts of the technologies remains of paramount importance".

[4] Cheng, Hui-Chan, *et al.* (1991): "Comparison of Performance and Attitude in Traditional and Computer Conferencing Classes" in *American Journal of Distance Education*, 53(3):51-64; ERIC Acc. No. EJ438088.

[5]A suggested reference is Standards for Evaluations of Educational Programs, Projects, and Materials, developed by the Joint Committee on Standards and Evaluation This reference lists all the ethical and procedural guides for evaluation and gives examples and suggestions.

[6] Shavelson, Richard J., *et al.* (1986): "Evaluating Student Outcomes from Telecourse Instruction. A Feasibility Study." (Sponsor: Corporation for Public Broadcasting, Washington, D.C.) ERIC Acc. No. ED311867

[7] Shavelson, Richard J., et al. (1986)

[8] UWM Department of Continuing and Vocational Education (1987): *Evaluation of Teaching/Learning at a Distance. Report of the Annual Conference on Teaching at a Distance.* (Third; Madison, Wisconsin, August 3-6, 1987). Volume II: Informational Session Papers and Abstracts. ERIC Acc. No. ED307850.

[9] Shavelson, Richard J., *et al.* (1986)

[10]See instrument samples in appendix C.

[11]Based upon sample mean subtracted from the true mean of the population mean divided by the standard deviation of the sample

[12]Craft, Elizabeth H. & Wagner, Norman (1988): "Instructional Television Fixed Service in Arizona" in *American Journal of Distance Education*, 2(1):76-80; ERIC Acc. No. EJ372451.

[13] Morehouse, Diane (1987): "Evaluating Interactive Television Methods, Findings, Issues" in *Evaluation of Teaching/Learning at a Distance. Report of the Annual Conference on Teaching at a Distance.* (Third; Madison, Wisconsin, August 3-6, 1987; Volume II: Informational Session Papers and Abstracts. ERIC Acc. No. ED307850.) pp. 119-122.

[14] Seamons, Alan (1987): "The Influence of Teaching Style and Instructional Device Use on Student Satisfaction and Preference" in *Evaluation of Teaching/Learning at a Distance. Report of the Annual Conference on Teaching at a Distance.* (Third; Madison, Wisconsin, August 3-6, 1987; Volume II: Informational Session Papers and Abstracts. ERIC Acc. No. ED307850.) pp. 119-122.

[15]Style was identified with an instrument called a "Gregoric Style Delineator," satisfaction using a Likert based evaluation form.

[16] Beare, Paul L. (1989): "The Comparative Effectiveness of Videotape, Audiotape, and Telelecture in Delivering Continuing Teacher Education" in *American Journal of Distance Education*, 1989, 3(2):57-66; ERIC Acc. No. EJ401019.

[17] Feasley, Charles E. (1987): "Evaluation of Student Outcomes" in *Evaluation of Teaching/Learning at a Distance. Report of the Annual Conference on Teaching at a Distance.* (Third; Madison, Wisconsin, August 3-6, 1987). Volume II: Informational Session Papers and Abstracts. ERIC Acc. No. ED307850.

[18] Jackman, Diane H. & Swan, Michael K. (1996): "Instructional Models Effective in Distance Education." Paper presented at the Annual Meeting of the American Educational Research Association (New York, NY, April 8-12, 1996); ERIC Acc. No. ED398867. The report does not specifically state its definition of a distance education classroom but by inference it refers to delivery of instruction via interactive television. Researchers often assume that their definition of distance education is the only definitive one!

[19] Biner, Paul M. et al. (1994): "Factors Underlying Distance Learner Satisfaction with Televised College-Level Courses" in *American Journal of Distance Education*, 8(1):60-71; ERIC Acc. No. EJ483695.

[20] see questionnaire appendix C

[21] Feasley, Charles E. (1987)

[22] Zhu, Erping (1996): "Meaning Negotiation, Knowledge Construction, and Mentoring in a Distance Learning Course" in *Proceedings of Selected Research and Development. Presentations at the 1996 National Convention of the Association for Educational Communications and Technology (18th, Indianapolis, IN. 1996* ERIC Acc No. ED 397 849. This study references the work of L.S. Vygotsky, *Mind in Society: The Development of Higher Psychological Processes* (Cambridge, MA: Harvard University Press, 1978). Also Tudge (1990). Vygotsky, The Zone Of Proximal Development And Peer Collaboration: Implications For Classroom Practice." In Moll, L.C. (Ed.), *Vygotsky And Education: Instructional Implications And Applications Of Sociohistorical Psychology.* (New York: Cambridge University Press, (no date supplied).

[23] Cheng, Hui-Chan, *et al.* (1991)

[24]Beare, Paul L. (1989) By comparison, the RAND/CPB report suggests that self selection is valuable and should not be eliminated.

[25] Whetzel, Deborah L., et al. (1996): "A Real World Comparison of the Effectiveness of Satellite Training and Classroom Training." In *Educational Research and Development*, 44(3):5-18; ERIC Acc. No. EJ532851.

[26]Models used were those developed by Joyce, Weil and Showers (1992) Models of Teaching Boston: Allyn and Bacon.

[27] Feasley, Charles E. (1987)

[28] Shavelson, Richard J., *et al* (1986)

[29] Felleni, Robert, et al. (1987): "Variables Affecting Adult Learning in a Distance Education Setting" and Morehouse, Diane (1987): "Evaluating Interactive Television: Methods, Findings and Issues" both in *Evaluation of Teaching/Learning at a Distance. Report of the Annual*

Conference on Teaching at a Distance. (Third; Madison, Wisconsin, August 3-6, 1987). Volume II: Informational Session Papers and Abstracts. ERIC Acc. No. ED307850.

[30] Hoffman, Jeff (1997): "The Learner Interaction Model for the Design of Interactive Television Courses" in *ED Journal*, April 1997, 11(4):J9-J12.

[31] Cheng, Hui-Chan, *et al.* (1991)

[32] Wagner, Ellen D. (1993): "Evaluating Distance Learning Projects: An Approach for Cross-Project Comparisons." Presented at the Annual Meeting of the Association for Educational Communications and Technology (15 ; New Orleans, LA: 1993); ERIC Clearinghouse IR530327; 32 pp. This is not a research document.

[33] In Saba, Farhad (Ed.): *Defining Concepts in Distance Education* (Madison, Wisconsin: Magna Publications, 1997).

[34] Hackman, Michael Z. and Walker, Kim B. (1994): "Perceptions of Proximate and Distance Learners Enrolled in University-Level Communication Courses: A Significant Non-Significant Finding." Paper presented at the Annual Meeting of the International Communication Association (44th; Sydney, New South Wales, Australia; July 11-15, 1994); ERIC Acc. No. ED372428.

[35] Egan, M. Winston and McCleary, Iva Dene (1989): "Program Design and Evaluation: Two-Way Interactive Television" in *American Journal of Distance Education*, 3(1):50-60; ERIC Acc. No. EJ392473. Design features included visual material, organization, and feedback.

[36] Kendall, Janet Ross and Oaks, Muriel (1992): "Evaluation of Perceived Teaching Effectiveness: Course Delivery via Interactive Video Technology versus Traditional Classroom Methods" in *Journal of Continuing Higher Education*, 40(3):2-12, Fall 1992; ERIC Acc. No. EJ451033.

[37] Shavelson, Richard J., et al. (1986).

[38] Shavelson, Richard J., et al. (1986) p.24

[39] WCTE is a section of WICHE, the Western Interstate Commission for Higher Education, Boulder, CO.

[40] Re: Wagner (1993) Reader Shoshanna Porter writes: " This is an excellent overview especially for schools that believe that existing courses can be easily transferred a distance ed. format and still provide a quality learning experience. This article brings out the differences that all aspects of quality preparation plays in the instructional environment."

[41] This work continues and is documented on the WWW through "The Flashlight Project" http://wiche.edu/flshlght/flashrpts.htm

[42] Willis, Barry (1992): "Making Distance Learning Effective: Key Roles and Responsibilities" in *Educational Technology*, 32(6):35-37, June 1992; ERIC Acc. No. EJ447500.

[43] Hardy, Darcy Walsh and Olcott, Donald, Jr. (1995): "Audio Teleconferencing and the Adult Learner: Strategies for Effective Practice" in *American Journal of Distance Education*, 9(1):44-60; ERIC Acc. No. EJ503484. This is not really a research report but is a useful document for planning any distance education experience. This is not really a research report but is a useful document for planning any distance education experience.

[44] Craft and Wagner (1988)

[45] Seamons (1987)

[46] Rowntree (1981): "Teaching Through Self Instruction" (p. 23); and Estabrooke (1981): "Evaluation and Grading of Student Work" both in Lambert, Michael P. & Welch, Sally R. (Eds.) *Home Study. Student Services Handbook* (Washington, D.C. National Home Study Council, 1981). Feedback comments.

[47] Hackman, Michael Z. and Walker, Kim B. (1994)

[48] Cheng, *et al.* (1991)

[49] Cited in Egan and McCleary (1989).